

# Using Logical Observation Identifier Names and Codes (LOINC) to Exchange Laboratory Data Among Three Academic Hospitals.

David M. Baorto, MD, PhD<sup>2,1</sup>, James J. Cimino, MD<sup>3</sup>, Curtis A. Parvin, PhD<sup>2</sup>, and Michael G. Kahn, MD, PhD<sup>1</sup>

<sup>1</sup>Section of Medical Informatics, Washington University, St. Louis, MO

<sup>2</sup>Division of Laboratory Medicine, Washington University, St. Louis, MO

<sup>3</sup>Department of Medical Informatics, Columbia University, New York, NY

*Using a standard set of names and codes to exchange electronic laboratory data would facilitate multiinstitutional research and data pooling. This need has led to the development of the Logical Observation Identifier Names and Codes (LOINC) database and its test naming convention. We conducted a study which required 3 academic hospitals (in 2 separate medical centers) to extract raw laboratory data from their local information system for a defined patient population, translate tests into LOINC, and provide aggregate data which could then be used to compare laboratory utilization. We found that the coding of local tests into LOINC can often be complex, especially the "Kind of Property" field, and apparently trivial differences in choices made by individual institutions can result in nonmatches in electronically pooled data. In our study, 72 - 86% of the failures of LOINC to match the same tests between different institutions were due to differences in local coding choices. LOINC has tremendous potential to eliminate the need for detailed human inspection during the pooling of laboratory data from diverse sites, and perhaps even a built-in capability to adjust matching stringency by selecting subsets of LOINC fields required to match. However, a quality, standard coding procedure at all sites is critical.*

## INTRODUCTION

Advocates of electronic medical record systems often propose the aggregation of patient data from multiple clinical databases to support multi-institutional or national health services research studies [1]. To accomplish this goal however, comparable clinical data from multiple institutions must be combined. Central to combining clinical information from disparate and independent sources is the use of a standardized set of terms or codes in which contributing organizations agree to encode their data. Once translated into an agreed-upon standard, data from different sources can be combined and analyzed.

In laboratory medicine, an effort to create a standard set of test names has achieved substantial momentum [2]. The Logical Observation Identifier Names and Codes (LOINC) database was initially motivated by the need to share laboratory test results among disparate clinical systems. The laboratory LOINC database has grown to over 6500 clinical test names or identifiers (version 1.0h). Recently a set of clinical terms, called the Clinical LOINC database, has been added (version 1.0i). LOINC databases are freely available via the Internet (<http://www.mcis.duke.edu/standards/termcode/loinc.htm>).

LOINC testnames are ASCII strings constructed by combining six component fields separated by a field delimiter (the colon character) [3]. Each name is assigned a unique code; the assigned code has no embedded semantics or interpretation. Institutions seeking to exchange laboratory data using the LOINC vocabulary must provide a mapping between each institutional-specific laboratory code or name in their system's term dictionary to LOINC names or codes. Each institution is responsible for representing their tests accurately in the LOINC vocabulary. Thus, the ability to use LOINC as a method to standardize laboratory test names from disparate institutions rests not only on the specificity and completeness of the LOINC vocabulary, but also on the ability of participating departments or institutions to encode their local tests correctly into LOINC identifiers.

We describe the results of an experiment which required combining laboratory test names and results from 3 hospitals at two independent academic institutions for the purpose of comparing laboratory test utilization in a specific clinical condition (congestive heart failure). Each institution was required to design and execute queries to extract raw data from their local laboratory information system, to provide their own mapping of local test names into LOINC identifiers, and to generate aggregate data for comparative analysis. Electronic correspondence between the two research groups defined the patient characteristics for inclusion in the study population and computations for constructing

aggregate comparative data, but no attempt was made to restrict or define the process of mapping local codes into LOINC identifiers. Investigators used the same information provided in the same version of the LOINC user's guide [3]. We focus here only on the difficulties with the encoding process for sharing data between two independent institutions. An analysis of the laboratory test utilization results appears elsewhere.

## METHODS

### Data Sources

We assessed the utility of LOINC for sharing laboratory data between 2 academic medical centers, Washington University School of Medicine in St. Louis, and Columbia University in New York (CPMC) by asking the following simple question: How does laboratory test utilization differ among these hospitals for all patients admitted between January 1, 1995 and December 31, 1995 with a primary discharge ICD-9 diagnosis of congestive heart failure (428.0)? During the study period, Washington University consisted of 2 different academic teaching hospitals: Barnes Hospital and Jewish Hospital (During 1996, these hospitals were merged as Barnes-Jewish Hospital).

### Database Queries

Each site was responsible for querying the local laboratory database. At Barnes Hospital, all patient and laboratory data are stored on an IBM mainframe computer in DB2 tables for a period of approximately 2 years. EASYTRIEVE queries produced ASCII output files containing 1) a listing of patient registration numbers, admission and discharge dates, for all patients admitted between January 1, 1995 and December 31, 1995 who had the primary diagnosis of congestive heart failure (ICD9 code 428.0) upon discharge and 2) all laboratory procedures reported on those patients including test identifier, battery identifier (how test was ordered), test result, and test date and time. The files were uploaded into Stata, a commercial statistical software package. For the present study, tests were listed in decreasing order of test volume for each institution. The most frequent 50 tests at each institution were used in subsequent analysis. The list of the most frequent 50 tests at Jewish Hospital were compiled from a 4 month period from September 1, 1995 to December 31, 1995, the period for which electronic data was available. Similarly, summaries from the CPMC database were obtained in two steps. In the first step, the financial database (IBM's PM/PA) was queried to identify all admissions with the same admission date range and ICD9-CM primary diagnosis as was used for the Barnes and Jewish

Hospital queries. For each admission, the medical record number, admission date and discharge date were obtained. A second query was then made to the clinical database (DB2) to obtain, for each medical record number, all laboratory data falling between the admission and discharge dates. "True Matches" were clearly the same laboratory test as determined by one of the authors who is a laboratory medicine physician.

### LOINC Coding

Each site was responsible for mapping local test names into LOINC names with no prior discussion of coding rules, except that LOINC version 1.0g was used. Each site then provided a listing in LOINC of the most common 50 tests ordered on the patient population described above. In practice, the LOINC coding at Barnes and Jewish Hospitals was done by the same person, while the coding at CPMC was done independently. Codes were also included for each test where a pre-existing code existed in the LOINC database (version 1.0g). Test names and summary data for each test were provided as follows: <local test identifier> | <LOINC code (if it existed in database)> | <Analyte> : <Kind of property measured> : <Time aspect> : <Sample type> : <Precision> | <#patients having test> | <total # ordered>. The fully specified LOINC name is between the second and third "|".

## RESULTS

### Reasons for nonmatches with the preexisting LOINC database.

A large proportion of the LOINC names for the top 50 tests at the three institutions did not have a pre-existing LOINC code in the LOINC database (version 1.0g). Among the top 50 tests at Barnes Hospital 31 tests had a matching LOINC code, at CPMC 16 tests had a matching LOINC code, and at Jewish Hospital 29 tests had a matching LOINC code. We further examined why many fully specified LOINC names among the 50 most frequent tests on CHF patients (19 at Barnes, 21 at Jewish, and 34 at CPMC) did not precisely match test names in the LOINC database.

Of the 19 testnames at Barnes that did not match to a LOINC code, 11 failed based on "sample type" and 10 of those failed because the tests at Barnes were done on "plasma" and LOINC only had a code for "serum", not "plasma/serum" or "plasma". Of the remaining 8, 4 failed based on "kind of property measured". In one case, the existing LOINC code for creatine kinase.MB is for catalytic activity, not mass concentration, which is measured at Barnes. The

Table 1. Matching Between Hospital Pairs Using Different LOINC Field Combinations

Hospital Pair	# of Matches out of 50 Between Each Hospital Pair Using the Following LOINC Fields					True Matches
	LOINC Code only	4 LOINC Fields	3 LOINC Fields	2 LOINC Fields	1 LOINC Field	
Barnes- Jewish	26	35	35	35	41	42
Barnes- CPMC	4	11	11	13	19	33
CPMC- Jewish	1	4	4	4	17	29

other 3 nonmatches based on "kind of property" were due to our assigning properties to those 3 tests which differed from properties assigned by LOINC for the identical test. For example, LOINC assigned the property "number fraction" for hematocrit while we assigned "volume fraction", we assigned the property "range" to erythrocyte size distribution while LOINC assigned "length", and we assigned "mass per entity" (entity being erythrocyte) for mean corpuscular hemoglobin while LOINC assigned "mass concentration".

Two testnames in the Barnes Hospital top 50 list were not present in the LOINC database, including "Carbon dioxide.calculated" and "Percent neutrophils". The final 2 nonmatches between Barnes and LOINC were due to a distinction in coagulation testing methodology that LOINC did not consider. The LOINC database codes one prothrombin time (PT) test called "coagulation tissue factor induced^^^ Patient: Time: PT: PPP" (see LOINC manual for explanation of fields). However, Barnes Hospital has 2 distinct prothrombin time tests which differ by the ISI number of the thromboplastin, one of the reagents used in the assay. While the fully specified LOINC name for these 2 tests would be identical and both actually match the testname in the LOINC database, these tests can yield vastly different results and should be distinguished.

The 21 nonmatches between the Jewish Hospital top 50 list in CHF patients and the LOINC database were due to the following reasons. 5 tests were components of blood gas analysis which would have matched except that the laboratory does not have distinct testnames for arterial and venous blood, and the LOINC database does not include test codes for ABG specimens of ambiguous origin. An additional 4 nonmatches were based on "sample type", and 4 were based on "kind of property measured". Seven of the top 50 Jewish Hospital testnames were missing from the LOINC database, including a number of Jewish Hospital blood gas components such as "Tidal Volume" and "Ventilation Mode".

The remaining 1 was the prothrombin time as discussed above.

Most of the nonmatches between the Columbia top 50 list and the LOINC database were based on "sample type", with 17 due to the plasma/serum issue mentioned previously. Three were due to "kind of property", in 2 cases of which CPMC reports a mass concentration, where the LOINC database only supported molar concentration. Other nonmatches were due to analyte choices. For example the CPMC coder chose to assign the analyte "erythrocytes" instead of "hematocrit".

#### Effect of lowering the stringency for matching among the 3 hospitals.

We examined how the top 50 lists at the 3 hospitals compared to each other. The number of true matches as well as the number of matches obtained by using different fields of the LOINC testname appear in Table 1. There were the fewest number of matches between the hospitals when the LOINC code was used as the only matching criterion. This was not surprising since many LOINC names at each institution had no associated LOINC code. Using the following 4 fields of the LOINC testname resulted in more matches between each pairwise combination of hospitals than the LOINC code alone: <analyte>:<sample type>:<kind of property>:<time aspect>. The <precision> field was not included because it was the same ("QN", or quantitative) for all tests in our lists. By sequentially removing the last LOINC field from the matching criteria, the matching stringency was reduced in a stepwise manner. Removing <time aspect> had no effect on the number of LOINC matches (transition from 4 to 3 LOINC fields). This is expected since all the tests in the top 50 lists were single timepoint measurements, not rates. Removing <kind of property> from the matching criteria also had little effect on the number of matches. Removing <Sample type> from the matching criteria (transition from 2 to 1 LOINC field), had the largest effect on increasing the number

of matches between the hospital pairs, although this severely impairs clinical relevance.

#### LOINC matching failures are primarily due to local coding choices.

It is notable that for all 3 hospital pairs, the number of true matches is greater than that which could be obtained by LOINC matching even at the lowest stringency. In Table 2, the total number of LOINC failures between each hospital pair is displayed. This represents the difference between the number of true matches, and the number of matches obtained by the 4 field LOINC match described above. We examined the reasons why LOINC failed to match in these cases, and divided them into 2 groups:

- 1) Failure due to local coding choices
- 2) Failure due to another laboratory factor

It is notable that, overall, 72% of LOINC matching failures are due to local coding choices. Between either Jewish or Barnes Hospital and CPMC, 72 - 86% of the LOINC failures are due to coding, while between Barnes and Jewish Hospitals, only 29% of the LOINC failures are due to coding. This is not surprising since the same person coded Barnes and Jewish Hospitals, while CPMC was coded independently. Some examples of LOINC matching failures due to local coding choices are: Analyte: CREATINE KINASE.TOTAL vs. CREATINE KINASE , Sample type: SERUM/PLASMA vs. PLASMA (for analytes where serum and plasma would be the same), and Property: PRES vs. PPRES (for partial pressure). An example of LOINC matching failure due to other laboratory factors is that the laboratory information system at one hospital does not store whether a blood gas analysis specimen is arterial or venous.

Table 2. Analysis of LOINC Matching Failures

Hospital Pair	Total LOINC failures	Failures Due to Coding	Failures Due to Laboratory
Barnes-Jewish	7	2	5
Barnes-CPMC	22	19	3
CPMC-Jewish	25	18	7

## DISCUSSION

We found that, even for common tests, there can be major differences in how individual hospitals code laboratory tests into LOINC. Coding error in other arenas, such as hospital discharge abstracts using ICD-9 has been studied by many investigators [4-9]. Incorrect principal diagnosis coding errors between 18.5% and 42.8% have been observed [10]. For DRG

encoding, error rates of 14.7% to 20.8% have been reported [10]. Such large error rates are not surprising because of the large degree of subjective interpretation and domain-specific knowledge is required for encoding clinical diagnoses. The amount of disagreement found in our study was surprising because encoding laboratory test names initially appeared to require far less ambiguity than would encoding clinical diagnoses or DRGs. In practice this may not be true.

Disagreements among local experts can also cause differences in LOINC coding even for common identical tests. The power of LOINC to be highly specific also makes it complex, and correct translation to LOINC requires a high resolution of knowledge of laboratory testing, precisely what properties are being measured and on what entity and by what method for each test. We found that in practice, physicians with good understanding of ordering and interpreting laboratory tests at their local institution, frequently don't have the resolution of knowledge to successfully translate all tests to LOINC. In our small study, we found the "kind of property" field created significant disagreements, even with the LOINC database itself. For example, the standard automated hematocrit, while frequently done by automated cell counting is a calculated value that represents a "volume fraction" (VFRC). However, in LOINC version 1.0g, this hematocrit was called a "number fraction" (NFRC), probably because an initial cell count is done prior to the calculation. Mismatches between our hospitals for tests such as "erythrocyte mean corpuscular volume" (MCV) were caused by one hospital assigning "entity volume" (ENTVOL) to the kind of property field (the entity being "RBC" (red blood cell), or erythrocyte), while another hospital assigned simply "volume" (VOL). The LOINC database assigned ENTVOL to the kind of property field and RBC to the sample type field in this case. However, for a conceptually similar laboratory test, the "platelet mean volume", the LOINC database assigned ENTVOL to the kind of property field and BLOOD to the sample type field instead of the entity, PLATELET, which would have been consistent with their choice for the MCV. While such differences may seem trivial, they could, in practice, prevent common lab tests from being considered the same, which are actually identical between institutions.

The "sample type" field frequently caused mismatches because of the issue of serum vs. plasma. Most analytes yield very similar results on serum or plasma and should be coded as "SER/PLAS" even if a laboratory does primarily one or the other. Certain analytes, however, yield different results on serum or plasma (e.g., total

protein and phosphorus), and <sample type> should be encoded as "SER" or "PLAS". All hospitals in this study coded <sample type> for the most part based on what samples are handled by their laboratory, and not based on analyte properties. This is the primary reason why reducing matching stringency from 2 LOINC fields (<analyte>:<sample type>) (Table 1) to 1 LOINC field (<analyte>) in this study greatly increased the number of matches. While more recent versions of LOINC have begun to sort out the serum/plasma issue, the version used for this study had not yet done so.

Most of the inconsistencies and errors will likely be corrected as LOINC matures. The lack of preexisting codes in the LOINC database for many of our most common tests was surprising, but more recent LOINC versions provide much more coverage. One of the advantages of the LOINC approach is that coverage in the LOINC database is not always necessary, because tests can be given a fully - specified name using LOINC naming conventions. For some applications, matching by LOINC code may be too specific, and relaxing conditions required for a match by selecting subgroups of test name fields may be preferred.

Occasionally, a facility will not have distinct internal test codes for clinically significant specimen type differences. The lack of distinction of arterial from venous blood for blood gas testing in the Jewish Hospital information system was just one example of this. In general, such tests cannot be assigned a single LOINC code because they would lack the granularity of LOINC and encompass several distinct LOINC codes. Receiving systems must have an approach for handling incoming fully-specified LOINC names that are more general than those in the LOINC database or in their own system. Such situations would need to be addressed on a case-by case basis, and the solution would be dependent on the reason for the data-exchange.

The knowledge-based approach of LOINC allows laboratory tests, and more recently clinical results, to be explicitly defined by their code and/or name. Theoretically, this allows unambiguous pooling of data from diverse sites without the requirement for post-coding human inspection. Our study highlights that this goal can potentially be reached only if a careful, standard LOINC coding procedure is used at all sites, performed by individuals with significant domain-specific expertise and well-educated in the LOINC system.

## Acknowledgments

Dr Kahn is supported by grant U01-LM05845 and Dr. Cimino by grant LM05857 from the National Library of Medicine.

## References

- [1] Institute of Medicine, Committee on Improving the Patient Record. Dick RS, Steen EB (eds). The computer-based patient record: An essential technology for health care, National academy Press, Washington D.C., 1991.
- [2] Forrey AW, McDonald CJ, DeMoor G, Huff SM, Leavelle D, Leland D, Fiers T, Charles L, Griffin B, Stalling F, Tullis A, Hutchins K, Baenziger J. Logical observation identifier names and codes (LOINC) database: A public use set of codes and names for electronic reporting of clinical laboratory test results. Clin. Chem 42(1):81-90, 1996.
- [3] LOINC User's Guide. Available from <http://www.mcis.duke.edu/standards/termcode/loinc.htm>.
- [4] Demlo LK, Campbell PM, Brown SS. Reliability of information abstracted from patient's medical records. Med Care 1978;16:995-1005.
- [5] Corn RF. The sensitivity of prospective hospital reimbursement to errors in patient data. Inquiry 1981;18:351-60.
- [6] Cohen E, Bernier D, Tam S, Schimel D, Postel AH, Scheidt S, Stamm JB. Data quality and DRGs: An assessment of the reliability of federal beneficiary discharge data in selected Manhattan hospitals. J Community Health 1985;10:238-46.
- [7] Iezzoni LI, Burnside S, Sickles L, Moskowitz MA, Sawitz E, Levine PA. Coding of acute myocardial infarction: Clinical and policy implications. Ann Intern Med 1988;109:745-51.
- [8] Hsia DC, Krushat WM, Fagan AB, Tebbutt JA, Kusserow RP. Accuracy of diagnostic coding for Medicare patients under the prospect payment system. N Engl J Med 1988;318:352-5.
- [9] Hsia DC, Ahern CA, Ritchie PB, Moscoe LM, Krushat WM. Medicare reimbursement accuracy under the prospective payment system, 1985 to 1988. JAMA 1992;268:896-9.
- [10] Green J, Wintfeld N. How accurate are hospital discharge data for evaluating effectiveness of care? Med Care 1993;31:719-31.